What is claimed is:

1. A method for indexing feature vector data space comprising the step of:

(a) adaptively approximating feature vectors on the basis of statistical distribution of feature vector data in the feature vector data space.

2. The method of claim 1, wherein the step (a) further comprises the steps of:

(a-1) measuring the statistical distribution of the feature vector data in the feature vector data space;

(a-2) estimating marginal distribution of the feature vector data using the statistical distribution;

(a-3) dividing the estimated marginal distribution into a plurality of grids in which a probability of disposing the feature vector data in each grid is uniform; and

(a-4) indexing the feature vector data space using the divided grids.

3. The method of claim 2, further comprising prior to step (a-4), the step of updating the grids on the basis of a previous probability distribution function and an updated probability distribution function, when new data is received.

4. The method of claim 2, wherein step (a-4) further comprises indexing using vector approximation (VA) files.

5. The method of claim 2, wherein a number of the plurality of grids is determined by a number of bits assigned to the dimension.

6. The method of claim 2, wherein step (a-2) further comprises the steps of:

(a-2-1) defining a probability distribution function using a weighted sum of a predetermined distribution function; and

(a-2-2) obtaining an estimated probability distribution function by estimating predetermined parameters using the probability distribution function defined in the step (a-2-1).

7. The method of claim 6, wherein step (a-2-2) further comprises obtaining the estimated probability distribution function by estimating the predetermined parameters using all N predetermined data in each estimation, wherein N is a positive integer, on the basis of an expectation-maximization algorithm using the probability distribution function defined in the step (a-2-1).

8. The method of claim 6, wherein the predetermined distribution function is the Gaussian function.

9. The method of claim 6, wherein the probability distribution function of step (a-2-1) is a one-dimensional signal, p(x), wherein $p(x)=\sum_{j=1}^{N} p(x|j)P(j)$, and wherein $p(x|j)$ is defined as

$$p(x|\,\mathrm{j}\,)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x-\mu_j)^2}{2\sigma_j^2}\right\},$$

wherein coefficient P(j) is a mixing parameter that satisfies the following criterion $0 \le P(j) \le 1$ and $\sum_{j=1}^{M} P(j)=1$.

10. The method of claim 6, wherein the estimated probability distribution function of step (a-2-2) is obtained by finding Φj, j=1,...,M, which maximizes $\Phi(\Phi_1,\ldots,\Phi_M)=\prod_{l=0}^{N} P(v[l]|(\Phi_1,\ldots,\Phi_M))$, where parameters v[$l$], l=1,...,N, is a given data set.

11. The method of claim 10, wherein the estimated parameters of step (a-2-2) are updated according to the following

equations $$\mu_j^{t+1}=\frac{\sum_{l=1}^{N} p(j|v[l])^t\, v[l]}{\sum_{l=1}^{N} p(j|v[l])^t},$$

$$(\sigma_j^2)^{t+1}=\frac{\sum_{l=1}^{N} p(j|v[l])^t\,(v[l]-\mu_j^t)^2}{\sum_{l=1}^{N} p(j|v[l])^t},$$ and

$$P(j)^{t+1} = \frac{1}{N}\sum_{l=1}^{N} p(j|v[l])^{t}$$, wherein t is a positive integer representing a number of iterations.

12. The method of claim 11, wherein the estimated parameter set of step (a-2-2) using N data v[l] is given as $\{P(j)^{N}\mu_j^{N},(\sigma_j^2)^{N}\}$, and the updated parameter set for new data v[N+1], coming in, is calculating using the following equations:

$$\mu_j^{N+1} = \mu_j^{N} + \theta_j^{N+1}(v[N+1] - \mu_j^{N}),$$

$$(\sigma_j^2)^{N+1} = (\sigma_j^2)^{N} + \theta_j^{N+1}[(v[N+1] - \mu_j^{N})^2 - (\sigma_j^2)^{N}],$$

$$P(j)^{N+1} = P(j)^{N} + \frac{1}{N+1}(P(j|v[N+1]) - P(j)^{N})$$, and

$$(\theta_j^{N+1})^{-1} = \frac{P(j|v[N])}{P(j|v[N+1])}(\theta_j^{N})^{-1} + 1.$$

13. The method of claim 11, wherein the step (a-2-2) further comprises:

measuring a change of a probability distribution function which is defined as $\dfrac{\rho = \int(\hat{p}_{old}(x) - \hat{p}_{new}(x))^2 dx}{\int \hat{p}_{old}(x)^2 dx}$ for each dimension, wherein a

previous probability distribution function is $\hat{P}_{old}(x)$, and an updated probability distribution function is $\hat{P}_{new}(x)$; and

updating an approximation for the dimension if ρ is larger than a predetermined threshold value.

14. The method of claim 2, wherein step (a-3) further comprises dividing a probability distribution function into the plurality of grids to make areas covered by each grid equal, wherein the plurality of grids have boundary points defined by c[*l*], *l* = 0,...,2^b, where b is a number of bits allocated and wherein the boundary points satisfy a criterion, $\int_{c[l]}^{c[l+1]} \hat{p}(x)dx = \frac{1}{2^b}\int_{c[0]}^{c[2b]} \hat{p}(x)dx$, and wherein the estimated probability distribution function is $\hat{p}(x)$.